

# Andreas H. W. Küpper

530 Riverside Dr, Apt 5H, New York, NY 10027

(203) 435-8819 · [ahwkuepper@gmail.com](mailto:ahwkuepper@gmail.com) · [ahwkuepper.info](mailto:ahwkuepper.info) · [github.com/ahwkuepper](https://github.com/ahwkuepper)

## Profile

---

German astrophysicist with more than 10 years of experience in

- data mining/munging/analysis, machine learning, and statistical modeling,
- programming, scripting, and high-performance computing,
- leadership, advising, and communication.

## Relevant Professional Experience

---

<b>Columbia University, New York, NY, <i>Hubble Research Fellow</i></b>	Since 2013
<ul style="list-style-type: none"><li>• Measured the mass of the Milky Way by fitting <math>10^6</math> tidal stream models to data.</li><li>• Studied dark matter substructure by producing and analyzing a data set of <math>10^9</math> stars.</li></ul>	
<b>Yale University, New Haven, CT, <i>Research Fellow</i></b>	2013
<ul style="list-style-type: none"><li>• Developed a Bayesian framework in Python/C for statistical modeling of tidal streams.</li></ul>	
<b>Universität Bonn, Germany, <i>Postdoctoral Researcher</i></b>	2011 – 2013
<ul style="list-style-type: none"><li>• Invented a now widely used <a href="#">algorithm</a> for efficient statistical modeling of tidal streams.</li></ul>	
<b>Universität Bonn, Germany, <i>Graduate Student Researcher</i></b>	2007 – 2011
<ul style="list-style-type: none"><li>• Studied the formation of tidal streams with high-performance N-body simulations.</li><li>• Published a popular <a href="#">C/Fortran code</a> for generating star cluster models.</li></ul>	

## Education

---

<b>Universität Bonn, Germany, PhD in Astronomy, <i>summa cum laude</i></b>	2007 – 2011
<b>Penn State University, PA, Bootcamp on Astrostatistics with R</b>	June 2010
<b>Universität Bonn, Germany, Diplom in Physics (MSc equivalent)</b>	2001 – 2007

## Expertise

---

Data analysis, machine learning, and statistical modeling

- Bayesian inference modeling, Markov Chain Monte Carlo, maximum likelihood estimation,
- logistic/linear/non-linear regression, kernel density estimation, k-nearest neighbor algorithms,
- minimum spanning tree algorithms, linear/non-linear least-square fitting, KS testing,
- bootstrapping/jackknife resampling, principal component analysis, Gaussian mixture models

Programming, scripting, and high-performance computing

- Python (NumPy, SciPy, matplotlib, Scikit-learn, Pandas), C, R, Linux scripting,
- Excel, Octave, JavaScript, Fortran, OpenMP, MPI, CUDA

Leadership, advising, and communication

- organization/coordination of 10+ larger meetings and workshops, student/postdoc representative,
- advisor to 8 PhD/MSc students, 50+ presentations at conferences/public events, 20+ publications